Linear Algebra

Samira Hossein Ghorban
s.hosseinghorban@ipm.ir

Fall, 2021

Review: A useful lemma for self-adjoint matrices

Lemma

Let $A \in M_n(\mathbb{F})$ be a self-adjoint matrix. Then there is an invertible matrix $P \in M_n(\mathbb{F})$ such that $P^*P = I$ and

$$A = P \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} P^*,$$

where $\lambda_1, \ldots, \lambda_n$ are eigenvalues of A.

- Singular value decomposition (SVD) is a generalization of this where $A \in M_{mn}(\mathbb{F})$ does not have to be self-adjoint or even square. No restriction at all!

AA^* and A^*A

- Let $A \in M_{mn}(\mathbb{F})$.
- The matrices $AA^* \in M_m(\mathbb{F})$ and $A^*A \in M_n(\mathbb{F})$ are self-adjoint and their eigenvalues are non-negative real numbers.
- There are unitary matrices $U \in M_m(\mathbb{F})$ and $V \in M_n(\mathbb{F})$ and diagonal matrices $D_1 \in M_m(\mathbb{R})$ and $D_2 \in M_n(\mathbb{R})$ such that

$$AA^* = UD_1U^* \qquad A^*A = VD_2V^*.$$

- We show that

$$D_1 = \begin{bmatrix} \lambda_1 & & & & & \\ & \ddots & & & & \\ & & \lambda_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix} \quad D_2 = \begin{bmatrix} \lambda_1 & & & & & \\ & \ddots & & & & \\ & & \lambda_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix}$$

Non-zero eigenvalues for AA^* and A^*A

Lemma

If $A \in M_{mn}(\mathbb{F})$ and $B \in M_{nm}(\mathbb{F})$, then $x^n f_{AB}(x) = x^m f_{BA}(x)$.

Lemma

Matrices AA^* and A^*A share the same non-zero eigenvalues with the same algebraic multiplicities.

Sigular Values

- Suppose that $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_r > 0$ where $r = \operatorname{rank} A = \operatorname{rank} AA^* = \operatorname{rank} AA^*$.
- Let $\sigma_i = \sqrt{\lambda_i}$ for each $1 \leqslant i \leqslant r$, and consider the m by n matrix

$$\Sigma = \begin{bmatrix} \sigma_1 & & & & \\ & \sigma_2 & & & \\ & & \ddots & & \\ & & & \sigma_r & \\ & & & & \\ & & & & \end{bmatrix}$$

- We have

$$D_1 = \Sigma\Sigma^*, \qquad\qquad D_2 = \Sigma^*\Sigma.$$

A candidate for a decomposition of A

- $AA^* = UD_1U^*$ $\qquad\qquad$ $A^*A = VD_2V^*$.
- $D_1 = \Sigma\Sigma^*$ $\qquad\qquad$ $D_2 = \Sigma^*\Sigma$.
- So, we obtain

$$AA^* = UD_1U^* = U\Sigma\Sigma^*U^* = U\Sigma V^*V\Sigma^*U^* = \Big(U\Sigma V^*\Big)\Big(U\Sigma V^*\Big)^*,$$

$$A^*A = VD_2V^* = V\Sigma^*\Sigma V^* = V\Sigma^*U^*U\Sigma V^* = \Big(U\Sigma V^*\Big)^*\Big(U\Sigma V^*\Big).$$

- It suggests that

$$A = U\Sigma V^*.$$

SVD

- Write V as $\begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}$. For each $1 \leqslant j \leqslant r$,

$$A^* A v_j = \sigma_j^2 v_j.$$

- So,

$$AA^*(Av_j) = \sigma_j^2 A v_j.$$

- $\sigma_j^2 \neq 0$ implies that $Av_j \neq 0$, and consequently the unit eigenvector for eigenvalue σ_j^2 is $\frac{1}{\|Av_j\|} A v_j$.

- Write U as $\begin{bmatrix} u_1 & \cdots & u_m \end{bmatrix}$. So, $u_j = \frac{1}{\|Av_j\|} A v_j$.

- Also, $\|Av_j\|^2 = v_j^* A^* A v_j = \sigma_j^2 \|v_j\|^2 = \sigma_j^2$, so $\|Av_j\| = \sigma_j$.

SVD

- Hence $Av_j = \sigma_j u_j$ for each $1 \leqslant j \leqslant r$.
- Consequently,

$$\begin{aligned} AV &= A\begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} \\ &= \begin{bmatrix} Av_1 & \cdots & Av_r & 0 & \cdots & 0 \end{bmatrix} \\ &= \begin{bmatrix} \sigma_1 u_1 & \cdots & \sigma_r u_r & 0 & \cdots & 0 \end{bmatrix} \\ & \begin{bmatrix} u_1 & \cdots & u_r & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} \sigma_1 & & & & \\ & \sigma_2 & & & \\ & & \ddots & & \\ & & & \sigma_r & \\ & & & & \\ & & & & \end{bmatrix} \\ &= U\Sigma. \end{aligned}$$

SVD

This is "the " SVD

$$\begin{aligned} A &= U\Sigma V^* \\ &= \sigma_1 u_1 v_1^* + \cdots + \sigma_r u_r v_r^* \end{aligned}$$

Pseudo-inverse of A

Assume

$$\Sigma = \begin{bmatrix} \sigma_1 & & & & \\ & \sigma_2 & & & \\ & & \ddots & & \\ & & & \sigma_r & \\ & & & & \\ & & & & \end{bmatrix} \qquad \Sigma^\dagger = \begin{bmatrix} \frac{1}{\sigma_1} & & & & \\ & \frac{1}{\sigma_2} & & & \\ & & \ddots & & \\ & & & \frac{1}{\sigma_r} & \\ & & & & \\ & & & & \end{bmatrix}$$

The **pseudo-inverse** of A is defined as

$$A^\dagger = V\Sigma^\dagger U^*.$$

$A^\dagger$ fulfills the role of A^{-1}, "as far as possible."

For singular matrix A^*A

- Let $x^\dagger := V^*\Sigma^+ Ub = A^\dagger b$.
- This $x^\dagger$ is a solution for $A^*Ax = A^*b$:

$$\begin{aligned} A^*Ax^\dagger &= A^*AA^\dagger b \\ &= V\Sigma^*U^*U\Sigma V^*V\Sigma^\dagger U^*b \\ &= V\Sigma^*\Sigma\Sigma^\dagger U^*b \\ &= V\Sigma^*U^*b \\ &= A^*b. \end{aligned}$$

- Also, we show that $x^\dagger = V^*\Sigma^+ Ub = A^\dagger b \in \arg\min_x \|b - Ax\|$.

$A^\dagger b$ is the optimal solution of $Ax = b$.

- The column space of $A^\dagger = V\Sigma^\dagger U^*$ is the space generated by the first r columns of V.
- The null space of $A^*A = V\Sigma^*\Sigma V^*$ is equal to the space generated by the last $n - r$ columns of V.
- Since V is unitary,

$$C(A^\dagger) \perp N(A^*A).$$

- Assume that z is a solution of the equation system $A^*Az = A^*b$.
- We know that $A^*AA^\dagger b = A^*b$. So

$$A^*A(z - A^\dagger b) = 0 \Rightarrow z - A^\dagger b \in N(A^*A).$$

- Let $v = z - A^\dagger b$. Thus $z = v + A^\dagger b$ where $v \in N(A^*A)$ and $A^\dagger b \in C(A^\dagger)$.
- So, $\|z\|^2 = \|v\|^2 + \|A^\dagger b\|^2 \geqslant \|A^\dagger b\|^2$.

Thank You!